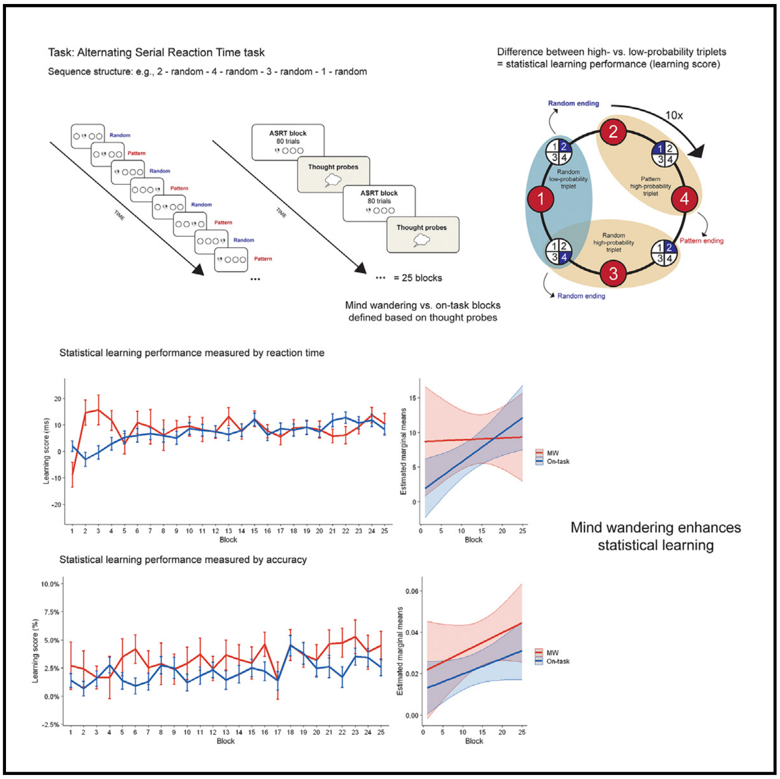
iScience

Article

Mind wandering enhances statistical learning

Graphical abstract



Highlights

- We spend much of our waking lives mind wandering (MW), but its purpose is unknown
- This study investigated the relationship between MW and statistical learning
- Participants performed a learning task while their focus was tracked
- Results show that MW enhances statistical learning, improving our predictive ability

Authors

Teodóra Vékony, Bence C. Farkas, Bianka Brezóczki, Matthias Mittner, Gábor Csifcsák, Péter Simor, Dezső Németh

Correspondence

simor.peter@ppk.elte.hu (P.S.),
dezso.nemeth@inserm.fr (D.N.)

In brief

Psychology

Vékony et al., 2025, iScience *28*, 111703
February 21, 2025
https://doi.org/10.1016/j.isci.2024.111703

CellPress

iScience

CellPress
OPEN ACCESS

Article

Mind wandering enhances statistical learning

Teodóra Vékony,[1,2] Bence C. Farkas,[3,4,5] Bianka Brezóczki,[6,7,8] Matthias Mittner,[9] Gábor Csifcsák,[9,13] Péter Simor,[7,10,11,13,*] and Dezső Németh[1,2,12,13,14,*]

[1]Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, INSERM, CRNS, Université Claude Bernard Lyon 1, 69500 Bron, France
[2]Gran Canaria Cognitive Research Center, Department of Education and Psychology, University of Atlántico Medio, 35017 Las Palmas de Gran Canaria, Spain
[3]UVSQ, INSERM, CESP, Université Paris-Saclay, 94807 Villejuif, France
[4]Institut du Psychotraumatisme de l'Enfant et de l'Adolescent, Conseil Départemental Yvelines et Hauts-de-Seine et Centre Hospitalier des Versailles, 78000 Versailles, France
[5]Centre de Recherche en Épidémiologie et en Santé des Populations, INSERM U1018, Université Paris-Saclay, Université Versailles Saint-Quentin, 94807 Paris, France
[6]Doctoral School of Psychology, Eötvös Loránd University, 1064 Budapest, Hungary
[7]Institute of Psychology, Eötvös Loránd University, 1064 Budapest, Hungary
[8]Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, 1117 Budapest, Hungary
[9]Department of Psychology, UiT The Arctic University of Norway, 9037 Tromsø, Norway
[10]Institute of Behavioral Sciences, Semmelweis University, 1085 Budapest, Hungary
[11]IMéRA Institute for Advanced Studies of Aix-Marseille University, 13004 Marseille, France
[12]BML-NAP Research Group, Institute of Psychology, Eötvös Loránd University & Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, 1071 Budapest, Hungary
[13]These authors contributed equally
[14]Lead contact
*Correspondence: simor.peter@ppk.elte.hu (P.S.), dezso.nemeth@inserm.fr (D.N.)
https://doi.org/10.1016/j.isci.2024.111703

SUMMARY

The human brain spends 30–50% of its waking hours engaged in mind-wandering (MW), a common phenomenon in which individuals either spontaneously or deliberately shift their attention away from external tasks to task-unrelated internal thoughts. Despite the significant amount of time dedicated to MW, its underlying reasons remain unexplained. Our pre-registered study investigates the potential adaptive aspects of MW, particularly its role in predictive processes measured by statistical learning. We simultaneously assessed visuomotor task performance as well as the capability to extract probabilistic information from the environment while assessing task focus (on-task vs. MW). We found that MW was associated with enhanced extraction of hidden, but predictable patterns. This finding suggests that MW may have functional relevance in human cognition by shaping behavior and predictive processes. Overall, our results highlight the importance of considering the adaptive aspects of MW, and its potential to enhance certain fundamental cognitive abilities.

INTRODUCTION

Mind wandering (MW) refers to a mental state when attention drifts away from the current task, becomes minimally constrained by the external environment, and descends into internally generated thoughts involving past experiences, imagined events, and anticipated future goals.[1–3] While in many everyday situations, we can dynamically adjust the focus of attention, and thus, disengagement from the external environment does not necessarily impact performance under low demands, the negative consequences of MW during various cognitive tasks have been extensively documented.[4,5] For instance, MW impairs reading comprehension,[6] sustained attention and executive control,[7,8] model-based decision-making,[9] explicit deterministic sequence learning,[10] working memory, and fluid intelligence.[11,12] On a behavioral level, reduced performance linked to MW is usually evidenced by worse accuracy, that is, failures to respond to targets, impulsive responses (e.g., quick responses to non-target items), or increased reaction time variability, all being indicative of suboptimal task-related cognitive control.[13–15] Moving beyond the behavioral domain, sensory decoupling during periods of MW was also consistently demonstrated by attenuated cortical responses that reflect the processing of specific stimuli in a task.[16]

Despite the apparent detriments of task-unrelated thoughts, the phenomenon of MW is ubiquitous, prompting researchers to investigate its potential advantages, such as its positive effects on planning or creative problem-solving.[4,15] For instance, influential theories suggest that MW may facilitate future planning and goal setting, by shifting the attention to personally relevant goals and concerns.[17] Accordingly, empirical studies propose that MW may serve as a self-reminding process for unresolved intentions.[18] A more recent study also provided evidence for the mnemonic benefits of MW in the domain of prospective memory. Spontaneous MW during an immersive virtual walk boosted the retrieval and execution of planned actions.[19] Although it is conceptually appealing to hypothesize about MW's adaptive significance, compelling and robust empirical evidence in this regard is still scarce and elusive to some extent.[20–23]

MW has been proposed to reflect transient offline states during wakefulness,[24] and interestingly, a growing body of research suggests that offline states are beneficial for learning and memory consolidation.[25–27] While previous studies on memory consolidation have focused on the retention of previously acquired information, recent studies suggest that memory consolidation processes are also critical for predictive processes.[28,29] Descriptive studies on the content of spontaneous thoughts indicate that MW is dominantly future-oriented, reflecting personally relevant goals, plans, and future behaviors.[30–32] These findings are consistent with the idea that one of the adaptive functions of self-generated thoughts may be to promote personally relevant mental simulations to anticipate and evaluate future scenarios.[2,4,15] However, the potential value of MW in predicting future outcomes has yet to be investigated. Considering that suboptimal cognitive control is linked to both heightened MW[33,34] and enhanced statistical learning[35,36]—an essential element of predictive processing[29,37]—it is plausible to hypothesize that MW could be associated with enhanced statistical learning. To address this gap, we present empirical evidence demonstrating that MW is associated with a beneficial impact on cognitive functioning, particularly in the realm of predictive processes.

In this study, we hypothesized that while sensory decoupling during MW would reduce task performance, it might simultaneously enhance the processing of environmental stimulus-outcome dependencies. Specifically, we predicted MW would be associated with improved statistical learning (see pre-registration: https://osf.io/cq6pg). This was tested using the Alternating Serial Reaction Time (ASRT) task, which measures visuomotor performance and implicit statistical learning.[38,39] Healthy adults (n = 135) completed 25 ASRT blocks. Thought probes assessed MW vs. on-task states (Figure 1). Results suggest that MW reduces response accuracy to external stimuli but improves the extraction of statistical patterns, leading to better predictions of future outcomes.

RESULTS

Thought probes

Out of the 135 participants, 117 reported MW at least once. Participants reported MW 30.13% of the time spent on the task (Figure 2). The median time participants first reported MW was Block 5 (M = 6.48; SD = 5.18). Of the time participants engaged in MW, 46.01% involved MB, and 23.80% of their focus was spontaneous rather than deliberate. Among the 117 participants, 21 reported MW only, without any mind blanking (MB), while 11 reported only MB with no MW content. Additionally, 34 participants reported only spontaneous MW, while 8 reported only deliberate MW without spontaneous instances (see results on MW/MB in Figures S1, S2, Tables S1, and S2, and results on spontaneous/deliberate focus in Figures S3, S4, Tables S3, and S4).

A simple linear regression was conducted to examine the effect of task Block (1–25) on mean MW scores. The regression equation was significant, $F(1, 3373) = 134.90$, $p < 0.001$, with an adjusted R^2 of 0.038. Block was a significant predictor of MW scores, $\beta = -0.026$, $t(3373) = -11.62$, $p < 0.001$, indicating that MW scores decreased (indicating more MW) as the task progressed (Figure 2). A simple linear regression was conducted to evaluate the relationship between Block (1–25) and the ratio of participants engaged in MW (0–100%). The regression equation was significant, $F(1, 23) = 74.95$, $p < 0.001$, with an adjusted R^2 0.76. The regression coefficient for Block was significant, $\beta = 0.010$, $t(23) = 8.66$, $p <. 001$, indicating that the ratio of the participants who engaged in MW increased as the task progressed.

Reaction times

With regards to RT, the model reported a main effect of Triplet Type, indicating faster RTs for high-probability trials than for low-probability trials ($b = -4.021$, 95% CI = [-4.607, −3.434], $F_{(1,\ 6362.38)} = 180.64$, $p < 0.001$) evidencing *statistical learning* during the task. The main effect of Block indicated decreasing RTs throughout the task ($b = -10.09$, 95% CI = [-12.661, −7.524], $F_{(1,\ 138.89)} = 60.37$, $p < 0.001$) evidencing altered *visuomotor performance* over time (Figure 3A). Statistical learning improved gradually during the task, as reflected by the interaction between Triplet Type and Block due to higher difference between high- and low-probability trials as the task progressed ($b = -0.816$, 95% CI = [-1.413, −0.218], $F_{(1,\ 6362.62)} = 7.16$, $p = 0.007$). Although the absence of an interaction between Triplet Type and MW indicated that the acquisition of statistical regularities reached a similar level during MW and on-task periods ($b = -0.494$, 95% CI = [-1.080, 0.093, $F_{(1,\ 6362.38)} = 2.73$, $p = 0.099$) (Figure 3B), the three-way interaction between Triplet Type, MW, and Block indicated higher RT difference between high- and low-probability triplets during MW compared to on-task periods at the beginning of the task ($b = 0.720$, 95% CI = [0.122, 1.317], $F_{(1,\ 6362.625)} = 5.58$, $p = 0.018$). This indicates that during MW the extraction of item probabilities was enhanced in initial blocks (Figure 3C). Full results can be found in Table S5.

Accuracy

Higher accuracy emerged for high-probability trials compared to low-probability trials (b = 0.014, 95% CI = [0.012, 0.016], $F_{(1,\ 178.68)} = 204.41$, $p < 0.001$), evidencing *statistical learning*. Participants showed higher overall accuracy (indicating better *visuomotor performance*) during on-task than during MW periods (b = −0.015, 95% CI = [-0.019, −0.012], $F_{(1,\ 117.32)} = 101.64$, $p < 0.001$) (Figure 4A). Accuracy decreased throughout the task (b = −0.011, 95% CI = [-0.014, −0.008], $F_{(1,\ 158.01)} = 55.57$, $p < 0.001$), with a more considerable drop in accuracy during MW periods with task progress (b = −0.005, 95% CI = [-0.007, −0.003], $F_{(1,3985.82)} = 19.21$, $p < 0.001$). Statistical learning improved as the task progressed (i.e., improving *statistical learning* with time) (b = 0.003, 95% CI = [0.001, 0.005], $F_{(1,\ 6326.31)} = 11.81$, $p < 0.001$). Most importantly, higher statistical learning emerged during MW periods (b = 0.003, 95% CI = [0.001, 0.005], $F_{(1,\ 2076.59)} = 9.38$, $p = 0.002$) (Figures 4B and 4C). That is, whereas visuomotor performance was worse during MW than on-task periods, the opposite was observed for statistical learning, showing

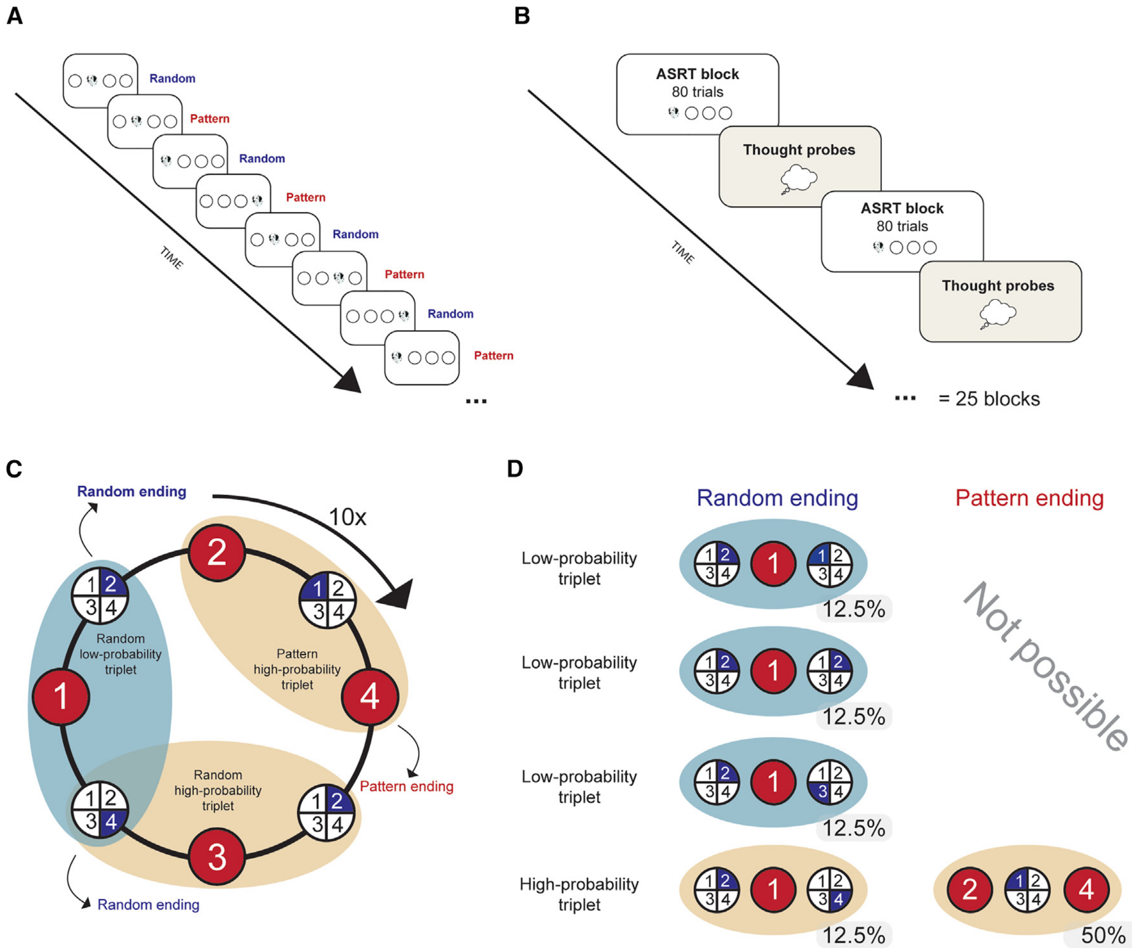


Figure 1. Experimental design and task structure of the ASRT task

(A) In the ASRT task, participants had to press keys corresponding to the location of the target stimulus (dog's head), where every second trial was part of an 8-element probabilistic sequence. Random elements were inserted among pattern elements to form the sequence.

(B) The experiment consisted of 25 blocks with thought probes administered after each block of 80 trials. Participants were asked to reflect on their thoughts and respond to three questions aimed at distinguishing between (1) on-task and MW (off-task) periods, (2) MW and mind blanking (MB) periods, and (3) deliberate vs. non-deliberate/spontaneous episodes.

(C) Formation of triplets in the task. Pattern elements are represented by red backgrounds (they constantly appear at that position throughout the task), and random elements are represented by blue backgrounds (they are always chosen from the four possible positions randomly). Every trial was categorized as the third element of three consecutive trials (a triplet). The probabilistic sequence structure resulted in a higher occurrence of some triplets (high-probability triplets) than others (low-probability triplets).

(D) The formation of high-probability triplets could have involved the occurrence of either two pattern trials and one random trial at the center, in 50% of trials, or two random trials and one pattern trial at the center (12.5% of trials). In total, 62.5% of all trials constituted the final element of a high-probability triplet, while the remaining 37.5% were the final elements of a low-probability triplet.

an advantage when participants engaged in MW. Full results can be found in Table S6. The described effects were also confirmed between subjects who engaged in varying amounts of MW; please refer to Tables S7–S10.

MW onset and learning performance

In order to further explore whether MW accelerates the statistical learning process, we examined how the timing of a participant's first MW episode (measured in task blocks) correlated with their learning ability in the initial phase of the task (average learning across the first five blocks). Our analysis revealed a significant negative correlation between the block number of the first MW report and early learning performance, as measured by RTs (RT: $r = -0.207$, $p = 0.025$; accuracy: $r = -0.048$, $p = 0.600$). This indicates that participants who reported MW earlier demonstrated a greater ability to differentiate between high- and low-probability trials in the initial stages of the task.

DISCUSSION

Humans can spend nearly 50% of their waking hours in MW, where attention shifts from external tasks to task-unrelated

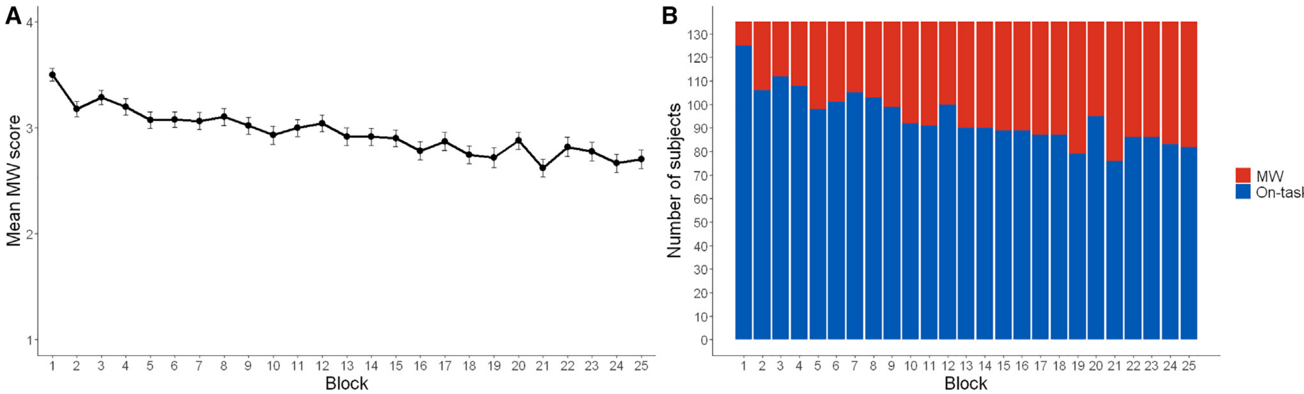


Figure 2. Change of MW over the course of the ASRT task

(A) Mean MW score per block, as reported by participants. The x axis represents the block number, and the y axis shows the average MW score (on a scale of 1–4, the lower score indicates more MW). The error bar indicates SEM.

(B) The number of participants engaged in MW in a given block. The x axis indicates the block number, while the y axis reflects the number of participants. Stacked bars differentiate between participants who reported MW (red) and those who reported on-task focus (blue). As the task progressed, both the overall MW and the number of participants reporting MW increased.

internal thoughts.[1,40,41] MW is known to reduce performance on cognitive tests, impact everyday activities[15,42–44] and have emotional cost.[1] Despite these drawbacks, the evolutionary benefits of MW are unclear. This study explores the potential benefits of MW, particularly its relationship to statistical learning. MW is thought to support divergent thinking/creativity and autobiographical planning/prospective simulations (for review see the study by Mooneyham et al.[15]) aligning with the association between MW and activity in the brain's default mode network[45,46] and the proposed functional relevance thereof.[2,47,48] However, recent studies have questioned MW's contribution to creative incubation[20–22] and its link to the default mode network.[49–51] Our study found that during MW, the extraction of predictable patterns in the environment is enhanced, resulting in improved statistical learning. Also, participants who experienced MW earlier were quicker to learn the underlying pattern in the task, supporting the hypothesis that MW might play a beneficial role in the early stages of statistical learning. Additionally, we observed changes in visuomotor performance, as evidenced by worse accuracy during MW (see also results on speed-accuracy tradeoff in Figure S5; Tables S11–S13 and on RT variability in Figure S6; Table S14, respectively). Given the importance of statistical learning and predictive processes in shaping behavior and neural computations,[52–54] these findings offer new insights into MW's role in cognitive functioning and everyday life.

Our findings of heightened statistical learning during MW align with the competition framework,[55,56] which posits an antagonistic relationship between cognitive control and statistical learning. Previous research indicates that statistical learning performance is negatively associated with control functions mediated by the prefrontal cortex.[35,57,58] For instance, EEG studies have shown that statistical learning negatively correlated with frontoparietal network activity.[59] Inhibitory non-invasive brain stimulation targeting the dorsolateral prefrontal cortex has improved predictive processing as measured by statistical learning.[60] On the other hand, MW has also been linked to the shift in the allocation of executive resources and impaired task-associated cognitive control. According to the executive failure view, MW episodes emerge as a result of the inability to maintain current goals via sustained task-focus and shielding against task-unrelated interference.[33] Studies have shown a negative association between MW and executive performance, such as the finger-tapping version of the classical random number generation task,[61] in which participants are asked to provide random sequences of finger taps to the rhythm of an ongoing metronome, while intermittently being probed about their mental states.[50,62,63] In line with our results, the aforementioned studies (along with many others,[3,13,14]) also showed increased behavioral variability during MW, providing support for the validity of our assessment of task-focus in the ASRT task (see Figure S6 and Table S14). Given that MW is coupled with impaired cognitive control, and statistical learning thrives when executive resources are depleted, it is likely that the successful extraction of the statistical contingencies during MW in our ASRT task was mediated by executive system failure. However, since we did not directly measure executive control (neither behaviorally, nor its neural correlates), future studies should address if MW benefits statistical learning through reduced executive control. This is especially relevant given the relatively low demands of the ASRT task, which may not have fostered MW at the expense of executive performance. Contextual factors, such as task difficulty, influence how individuals allocate attentional resources between task-focus and MW[64,65] and tasks that do not demand our full attention can lead to task-unrelated thoughts while maintaining optimal performance.

Studies on the neural correlates of MW suggest that impaired executive control is not necessarily the *sine qua non* of MW. Reduced amplitude of canonical event-related potentials (P100, N100, and P300) in EEG signals appeared to be robust markers of dampened cortical processing when participants' mind wandered, compared to periods when participants focused on the task, that in turn, elicited larger evoked potentials.[16] Interestingly, reduced cortical processing linked to

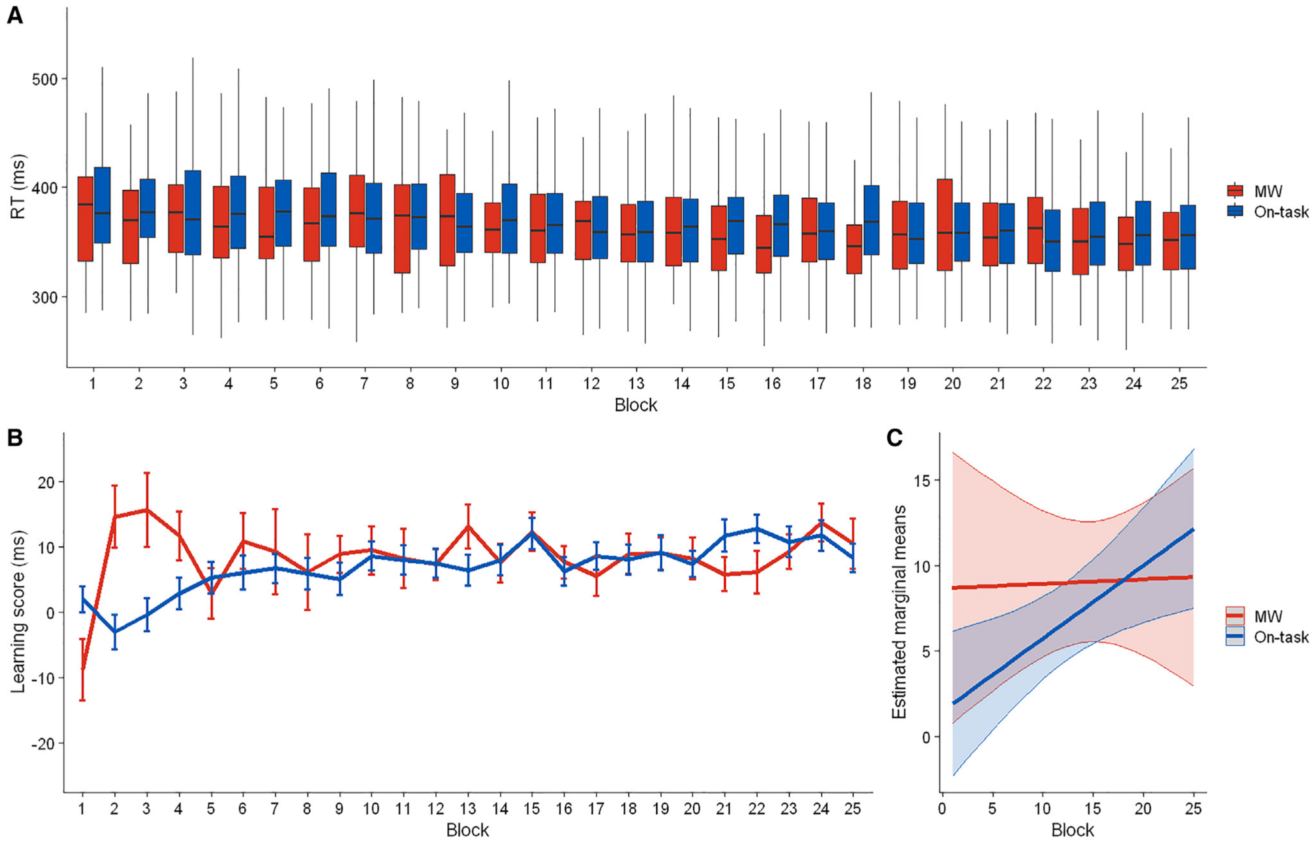


Figure 3. Visuomotor performance and statistical learning over the course of the task in the MW vs. on-task periods measured by reaction times

(A) Reaction time in ms plotted as a function of task progress. Red boxes indicate data from the MW (off-task) periods, and blue boxes that of the on-task periods. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), whiskers show 1.5 × IQR, and horizontal notches show the median.

(B) Raw statistical learning scores (difference between high- and low-probability trials) over the course of the task in the MW vs. on-task periods. The y axes indicate learning scores calculated and the x axes mark the blocks of the task. The red color indicates MW periods, and the blue color the on-task periods. Higher values represent better statistical learning (larger difference between high- and low-probability trials).

(C) Estimated marginal means of reaction time learning scores. Error bars and bands represent standard errors.

(B and C) Statistical learning was larger at the beginning of the task during MW periods (Triplet Type × MW × Block interaction).

See also Table S5.

task-unrelated thoughts was observed in response to both target and distractor stimuli, indicating that MW reflects a general decoupling from the external environment instead of failures in task-relevant processing and problems of distraction due to impaired executive control.[66] This argument is strengthened by the fact that the aforementioned P300 component has been linked to predictive processes—both with stimulus-locked and response-locked event-related potentials—measured with the same statistical learning task used in the current study.[67,68] Future studies directly testing the neural correlates of MW during statistical learning and predictive processes seem highly warranted.

Our findings suggest that MW, despite being linked to sensory decoupling, surprisingly facilitates processing of probabilistic sensory patterns. We propose that MW presents a transient, spatially localized offline state[24,69–71] that facilitates information processing during statistical learning. This may be explained by rapid memory consolidation of statistical information during sensory decoupling. The stabilization of memory traces is known to be either time-dependent or sleep-dependent,[25] with the latter being linked to low-frequency neural activity.[72,73] Although MW is associated with slow waves,[24,71] these slow waves are generated in resting wakefulness and expressed in more localized networks, a phenomenon also known as local sleep.[24,74,75] Our findings suggest that enhanced statistical learning observed during MW may be driven by memory consolidation associated with local sleep in the waking brain.[25,76] This implies the existence of a third category of memory consolidation—in addition to sleep- and time-dependent consolidation—referred to as local sleep-dependent consolidation. This type of consolidation may provide an explanation for the inconsistent findings of sleep-dependent memory consolidation in procedural or statistical learning tasks.[77] The brain consolidates learned material using local sleep waves during task-unrelated MW, fitting the

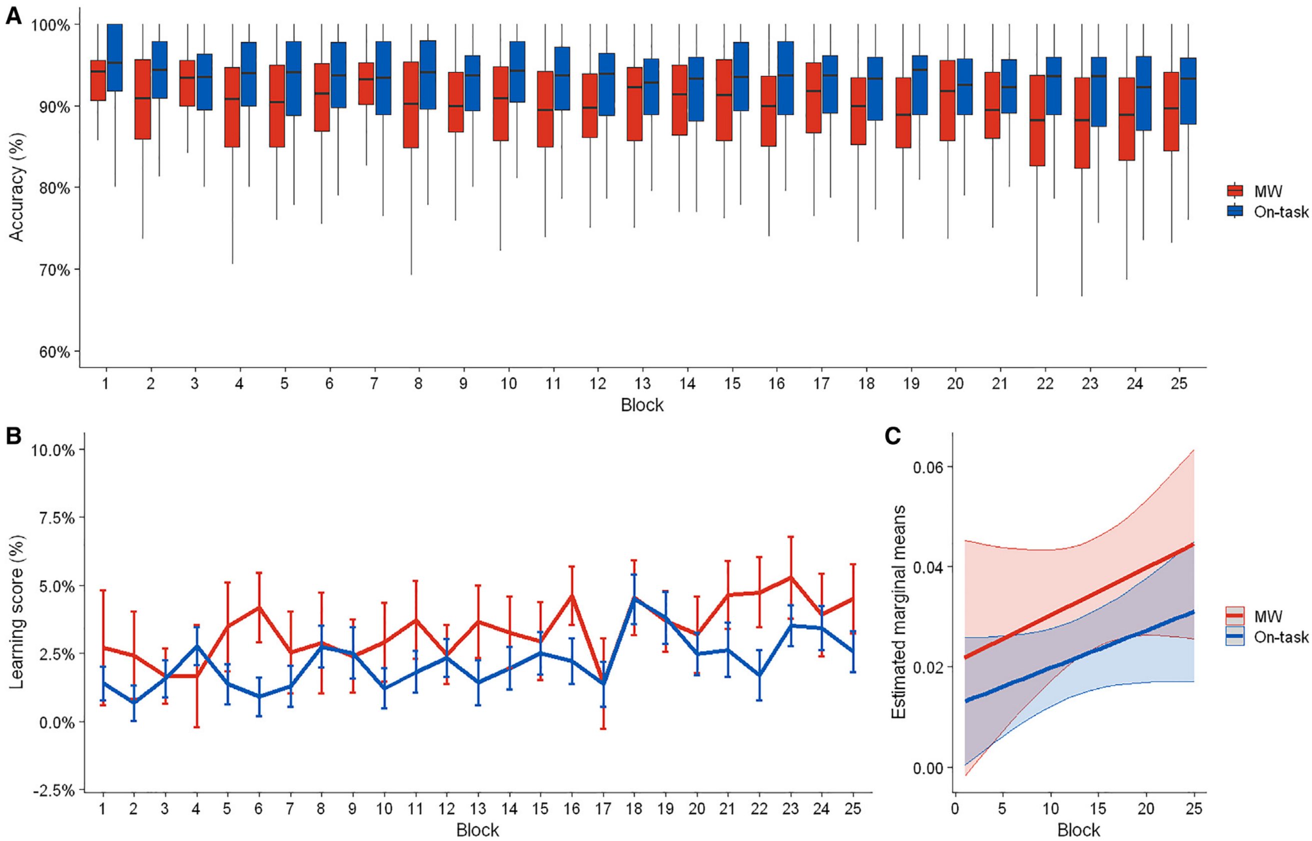


Figure 4. Visuomotor performance and statistical learning over the course of the task in the MW vs. on-task periods measured by accuracy
(A) Accuracy in percentage plotted as a function of task progress. Red boxes indicate data from the MW (off-task) periods, and blue boxes that of the on-task periods. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), whiskers show 1.5 × IQR, and horizontal notches show the median. During MW periods, participants were less accurate than during on-task periods.
(B) Raw statistical learning scores (difference between high- and low-probability trials) over the course of the task in the MW vs. on-task periods. The y axes indicate learning scores calculated and the *x* axes mark the blocks of the task. The red color indicates MW periods, and the blue color the on-task periods. Higher values represent better statistical learning (larger difference between high- and low-probability trials).
(C) Estimated marginal means of accuracy learning scores. Error bars and bands represent standard errors.
(B and C) Statistical learning was larger during MW periods (Triplet Type × MW interaction).
See also Table S6.

opportunistic theory of memory consolidation, which posits that consolidation occurs during awake, asleep, or local-sleep states.[27] In addition, our finding fits into the emerging number of studies indicating that even ultra-short periods of post-learning waking rest are beneficial for the stabilization of memory traces.[25,76] To test this hypothesis, it will be essential to conduct imaging, neurophysiological studies or experimental manipulations to induce sleep-like neural activity (i.e., increasing slow frequency oscillations in wakefulness) to provide direct evidence regarding the relationship between MW, local sleep, and enhanced learning performance. Future research should aim to clarify whether (a) MW directly facilitates predictive processing through local sleep-mediated memory consolidation by demonstrating a direct association between local slow waves and learning efficiency, or if (b) MW increases reliance on predictive processes indirectly, by consuming executive functions, which in turn improves consolidation of implicitly processed information.

Our assumptions imply that the positive association between MW and memory consolidation may be limited to undemanding, effortless task conditions, allowing the brain to enter into periods of waking rest. In contrast, task performance during MW may be disrupted only when facing higher cognitive load that requires executive control (but see results on mental effort in Figure S7; Tables S15 and S16). Therefore, our findings pointing to the benefits of MW may not generalize to such contexts. Accordingly, the beneficial influence of MW on cognitive performance was mainly observed under undemanding task conditions which allowed the mind to wander more frequently.[19,23] Nevertheless, given the paucity of data in this regard, the limits and boundary conditions of MW-related cognitive benefits are still a question that warrants future research.

Our findings might appear to contradict previous results where impaired sequence learning was previously reported during MW.[10] This previous study compared explicit and implicit conditions, finding that only explicit learning was impaired and implicit learning remained intact. It is crucial to recognize that deterministic and probabilistic sequence learning (investigated in our study) are fundamentally different. Probabilistic learning involves higher-order associations with varying level of uncertainty, so that predicting a current event requires considering the statistical constellations of more remote ones. This makes probabilistic learning unique in terms of relying more heavily on predictions in the face of the statistical interdependencies between immediate stimulus sequences, and consequentially, less prone to developing conscious awareness about task structure.[78] To gain a clearer understanding, future research needs to directly examine how MW influences both deterministic and probabilistic learning.

While here we argue that MW facilitates the extraction of statistical relationships in the environment, a process that commonly occurs without conscious awareness, some results on the influence of MW on memory encoding points at detrimental effects.[65] However, there is emerging evidence that MW may actually enhance encoding under certain conditions. Specifically, when the content of MW is related to the material being processed, it may enhance episodic memory encoding.[79] Additionally, studies have showed that stimulus-dependent thoughts enhance retrieval success in an incidental learning task, while stimulus-independent thoughts were not interfering with performance.[80,81] This distinction between task-relevant and task-irrelevant MW may explain the seemingly contradictory effects of MW on learning outcomes. Altogether, findings on the effects of MW on learning are mixed, and may depend on several factors, including the type of to-be-learned information, task demands, as well as the content of MW episodes. Future studies should aim to clarify the specific circumstances under which MW can exert a positive effect on memory encoding.

MW may hinder precise reactions to external stimuli, but it can also improve the recognition of consistent patterns, leading to better event prediction. Future studies could explore if these findings apply to other memory domains (e.g., episodic encoding of memory traces). MW might benefit the automatic acquisition of predictable patterns, which occurs without effortful processes. Statistical learning, crucial for predicting stimulus-outcome dependencies, aligns with MW's prospective nature, central to planning and future-oriented behavior. The extraction of environmental regularities might be fundamental to MW's role. Since statistical learning is vital for skill and habit development,[82–85] our results might generalize to these learning functions. Consequently, our findings could inform on MW's advantages in various learning domains, including language acquisition, motor skills, music learning, and social skill development.

Limitations of the study

This study lacked analysis stratified by age, sex, gender, ancestry, race, and ethnicity. This omission may restrict the generalizability of the findings to broader populations.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by Prof. Dezső Németh, dezso.nemeth@inserm.fr.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Behavioral data have been deposited at OSF and are publicly available at https://doi.org/10.17605/OSF.IO/T4AB3 as of the date of publication.
- All original code has been deposited at OSF and is publicly available at https://doi.org/10.17605/OSF.IO/T4AB3 as of the date of publication.
- The behavioral task can be previewed at https://app.gorilla.sc/openmaterials/581806. Any additional information required to reanalyze the data reported in this article is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the Chaire de Professeur Junior Program by Inserm and French National Grant Agency (ANR-22-CPJ1-0042-01); the National Brain Research Program project NAP2022-I-2/2022 (D.N.); Hungarian National Research, Development and Innovation Office grant NKFI FK 142945 (P.S.) and 128016 (D.N.); Janos Bolyai scholarship of the Hungarian Academy of Sciences (P.S.).

AUTHOR CONTRIBUTIONS

Conceptualization: T.V., M.M., G.C., P.S., and D.N.; methodology, T.V., B.C.F., B.B., M.M., G.C., P.S., and D.N.; software: T.V. and B.C.F.; validation: T.V., B.C.F., B.B., M.M., G.C., P.S., and D.N.; formal analysis: T.V., B.C.F., and B.B.; investigation: T.V. and D.N.; resources: T.V., B.B., M.M., G.C., P.S., and D.N.; data curation: T.V., B.B., and B.C.F.; writing—original draft: T.V., B.C.F., B.B., M.M., G.C., P.S., and D.N.; writing—review and editing: T.V., B.C.F., B.B., M.M., G.C., P.S., and D.N.; visualization: T.V. and B.C.F.; supervision: M.M., G.C., P.S., and D.N.; project administration: T.V., G.C., P.S., and D.N.; funding acquisition: P.S. and D.N.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT 3.5 in order to improve the writing style of our manuscript. The authors reject using AI for scientific content creation. However, the authors believe that it helps fostering the equality of native and non-native English speakers in order for them to have the same opportunities. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111703.

Received: July 19, 2024
Revised: October 14, 2024
Accepted: December 24, 2024
Published: January 4, 2025

REFERENCES

1. Killingsworth, M.A., and Gilbert, D.T. (2010). A wandering mind is an unhappy mind. Science *330*, 932. https://doi.org/10.1126/science.1192439.
2. Buckner, R.L., and Carroll, D.C. (2007). Self-projection and the brain. Trends Cogn. Sci. *11*, 49–57. https://doi.org/10.1016/j.tics.2006.11.004.
3. Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., and D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. Acta Psychol. *136*, 370–381. https://doi.org/10.1016/j.actpsy.2011.01.002.
4. Smallwood, J., and Schooler, J.W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. Annu. Rev. Psychol. *66*, 487–518. https://doi.org/10.1146/annurev-psych-010814-015331.
5. Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., and Sayette, M.A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn. Sci. *15*, 319–326. https://doi.org/10.1016/j.tics.2011.05.006.
6. McVay, J.C., and Kane, M.J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. J. Exp. Psychol. Gen. *141*, 302–320. https://doi.org/10.1037/a0025250.
7. Andrillon, T., Burns, A., Mackay, T., Windt, J., and Tsuchiya, N. (2021). Predicting lapses of attention with sleep-like slow waves. Nat. Commun. *12*, 3657. https://doi.org/10.1038/s41467-021-23890-7.
8. McVay, J.C., and Kane, M.J. (2012). Drifting from slow to "D'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. J. Exp. Psychol. Learn. Mem. Cogn. *38*, 525–549. https://doi.org/10.1037/a0025896.
9. Liu, S., Rabovsky, M., and Schad, D.J. (2023). Spontaneous mind wandering impairs model-based decision making. PLoS One *18*, e0279532. https://doi.org/10.1371/journal.pone.0279532.
10. Brosowsky, N.P., Murray, S., Schooler, J.W., and Seli, P. (2021). Attention need not always apply: Mind wandering impedes explicit but not implicit sequence learning. Cognition *209*, 104530. https://doi.org/10.1016/j.cognition.2020.104530.
11. Mrazek, M.D., Franklin, M.S., Phillips, D.T., Baird, B., and Schooler, J.W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. Psychol. Sci. *24*, 776–781. https://doi.org/10.1177/0956797612459659.
12. Mrazek, M.D., Smallwood, J., Franklin, M.S., Chin, J.M., Baird, B., and Schooler, J.W. (2012). The role of mind-wandering in measurements of general aptitude. J. Exp. Psychol. Gen. *141*, 788–798. https://doi.org/10.1037/a0027968.
13. Allan Cheyne, J., Solman, G.J.F., Carriere, J.S.A., and Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. Cognition *111*, 98–113. https://doi.org/10.1016/j.cognition.2008.12.009.
14. Bastian, M., and Sackur, J. (2013). Mind wandering at the fingertips: Automatic parsing of subjective states based on response time variability. Front. Psychol. *4*, 573. https://doi.org/10.3389/fpsyg.2013.00573.
15. Mooneyham, B.W., and Schooler, J.W. (2013). The costs and benefits of mind-wandering: A review. Can. J. Exp. Psychol. *67*, 11–18. https://doi.org/10.1037/a0031569.
16. Kam, J.W.Y., Rahnuma, T., Park, Y.E., and Hart, C.M. (2022). Electrophysiological markers of mind wandering: A systematic review. Neuroimage *258*, 119372. https://doi.org/10.1016/j.neuroimage.2022.119372.
17. Klinger, E., and Miles Cox, W. (2008). In Motivation and the Theory of Current Concerns. Handbook of Motivational Counseling: Concepts, Approaches, and Assessment (John Wiley & Sons Ltd), pp. 1–27. https://doi.org/10.1002/9780470713129.ch1.
18. Mason, M.F., and Reinholtz, N. (2015). Avenues down which a self-reminding mind can wander. Motiv. Sci. *1*, 1–21. https://doi.org/10.1037/mot0000011.
19. Girardeau, J.C., Ledru, R., Gaston-Bellegarde, A., Blondé, P., Sperduti, M., and Piolino, P. (2023). The benefits of mind wandering on a naturalistic prospective memory task. Sci. Rep. *13*, 11432. https://doi.org/10.1038/s41598-023-37996-z.
20. Smeekens, B.A., and Kane, M.J. (2016). Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. Psychol. Aesthet. Creat. Arts *10*, 389–415. https://doi.org/10.1037/aca0000046.
21. Steindorf, L., Hammerton, H.A., and Rummel, J. (2021). Mind wandering outside the box—About the role of off-task thoughts and their assessment during creative incubation. Psychol. Aesthet. Creat. Arts *15*, 584–595. https://doi.org/10.1037/aca0000373.
22. Murray, S., Liang, N., Brosowsky, N., and Seli, P. (2024). What are the benefits of mind wandering to creativity? Psychol. Aesthet. Creat. Arts *18*, 403–416. https://doi.org/10.1037/aca0000420.
23. Baird, B., Smallwood, J., Mrazek, M.D., Kam, J.W.Y., Franklin, M.S., and Schooler, J.W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. Psychol. Sci. *23*, 1117–1122. https://doi.org/10.1177/0956797612446024.
24. Andrillon, T., Windt, J., Silk, T., Drummond, S.P.A., Bellgrove, M.A., and Tsuchiya, N. (2019). Does the mind wander when the brain takes a break? Local sleep in wakefulness, attentional lapses and mind-wandering. Front. Neurosci. *13*, 949. https://doi.org/10.3389/fnins.2019.00949.
25. Wamsley, E.J. (2022). Offline memory consolidation during waking rest. Nat. Rev. Psychol. *1*, 441–453. https://doi.org/10.1038/s44159-022-00072-w.
26. Wamsley, E.J. (2019). Memory consolidation during waking rest. Trends Cogn. Sci. *23*, 171–173. https://doi.org/10.1016/J.TICS.2018.12.007.
27. Mednick, S.C., Cai, D.J., Shuman, T., Anagnostaras, S., and Wixted, J.T. (2011). An opportunistic theory of cellular and systems consolidation. Trends Neurosci. *34*, 504–514. https://doi.org/10.1016/J.TINS.2011.06.003.
28. Hauptmann, B., Reinhart, E., Brandt, S.A., and Karni, A. (2005). The predictive value of the leveling off of within-session performance for procedural memory consolidation. Brain Res. Cogn. Brain Res. *24*, 181–189. https://doi.org/10.1016/j.cogbrainres.2005.01.012.
29. Janacsek, K., and Nemeth, D. (2012). Predicting the future: From implicit learning to consolidation. Int. J. Psychophysiol. *83*, 213–221. https://doi.org/10.1016/j.ijpsycho.2011.11.012.
30. Baird, B., Smallwood, J., and Schooler, J.W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. Conscious. Cogn. *20*, 1604–1611. https://doi.org/10.1016/j.concog.2011.08.007.
31. Gross, M.E., Smith, A.P., Graveline, Y.M., Beaty, R.E., Schooler, J.W., and Seli, P. (2021). Comparing the phenomenological qualities of stimulus-independent thought, stimulus-dependent thought and dreams using experience sampling: Phenomenology of SIT, SDT, and dreams. Philos. Trans. R. Soc. Lond. B Biol. Sci. *376*, 20190694. https://doi.org/10.1098/rstb.2019.0694.
32. Baird, B., Aparicio, M.K., Alauddin, T., Riedner, B., Boly, M., and Tononi, G. (2022). Episodic thought distinguishes spontaneous cognition in waking from REM and NREM sleep. Conscious. Cogn. *97*, 103247. https://doi.org/10.1016/j.concog.2021.103247.

33. McVay, J.C., and Kane, M.J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). Psychol. Bull. *136*, 188–197. https://doi.org/10.1037/a0018298.

34. Smallwood, J., and Schooler, J.W. (2006). The restless mind. Psychol. Bull. *132*, 946–958. https://doi.org/10.1037/0033-2909.132.6.946.

35. Virag, M., Janacsek, K., Horvath, A., Bujdoso, Z., Fabo, D., and Nemeth, D. (2015). Competition between frontal lobe functions and implicit sequence learning: evidence from the long-term effects of alcohol. Exp. Brain Res. *233*, 2081–2089. https://doi.org/10.1007/s00221-015-4279-8.

36. Ambrus, G.G., Vékony, T., Janacsek, K., Trimborn, A.B.C., Kovács, G., and Nemeth, D. (2019). When less is more: enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. bioRxiv, 198515. https://doi.org/10.1101/198515.

37. Conway, C.M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. Neurosci. Biobehav. Rev. *112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032.

38. Howard, J.H., and Howard, D.V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. Psychol. Aging *12*, 634–656. https://doi.org/10.1037/0882-7974.12.4.634.

39. Nemeth, D., Janacsek, K., Londe, Z., Ullman, M.T., Howard, D.V., and Howard, J.H. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. Exp. Brain Res. *201*, 351–358. https://doi.org/10.1007/s00221-009-2024-x.

40. Maillet, D., Beaty, R.E., Jordano, M.L., Touron, D.R., Adnan, A., Silvia, P.J., Kwapil, T.R., Turner, G.R., Spreng, R.N., and Kane, M.J. (2018). Age-related differences in mind-wandering in daily life. Psychol. Aging *33*, 643–653. https://doi.org/10.1037/pag0000260.

41. Kane, M.J., Brown, L.H., McVay, J.C., Silvia, P.J., Myin-Germeys, I., and Kwapil, T.R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. Psychol. Sci. *18*, 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x.

42. McVay, J.C., Kane, M.J., and Kwapil, T.R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. Psychon. Bull. Rev. *16*, 857–863. https://doi.org/10.3758/PBR.16.5.857.

43. Szpunar, K.K., Moulton, S.T., and Schacter, D.L. (2013). Mind wandering and education: From the classroom to online learning. Front. Psychol. *4*, 495. https://doi.org/10.3389/fpsyg.2013.00495.

44. Yanko, M.R., and Spalek, T.M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. Hum. Factors *56*, 260–269. https://doi.org/10.1177/0018720813495280.

45. Christoff, K., Irving, Z.C., Fox, K.C.R., Spreng, R.N., and Andrews-Hanna, J.R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. Nat. Rev. Neurosci. *17*, 718–731. https://doi.org/10.1038/nrn.2016.113.

46. Mittner, M., Boekel, W., Tucker, A.M., Turner, B.M., Heathcote, A., and Forstmann, B.U. (2014). When the brain takes a break: A model-based analysis of mind wandering. J. Neurosci. *34*, 16286–16295. https://doi.org/10.1523/JNEUROSCI.2062-14.2014.

47. Shi, L., Sun, J., Xia, Y., Ren, Z., Chen, Q., Wei, D., Yang, W., and Qiu, J. (2018). Large-scale brain network connectivity underlying creativity in resting-state and task fMRI: Cooperation between default network and frontal-parietal network. Biol. Psychol. *135*, 102–111. https://doi.org/10.1016/j.biopsycho.2018.03.005.

48. Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., and Kawashima, R. (2011). Failing to deactivate: The association between brain activity during a working memory task and creativity. Neuroimage *55*, 681–687. https://doi.org/10.1016/j.neuroimage.2010.11.052.

49. Groot, J.M., Boayue, N.M., Csifcsák, G., Boekel, W., Huster, R., Forstmann, B.U., and Mittner, M. (2021). Probing the neural signature of mind wandering with simultaneous fMRI-EEG and pupillometry. Neuroimage *224*, 117412. https://doi.org/10.1016/j.neuroimage.2020.117412.

50. Groot, J.M., Csifcsák, G., Wientjes, S., Forstmann, B.U., and Mittner, M. (2022). Catching wandering minds with tapping fingers: neural and behavioral insights into task-unrelated cognition. Cereb. Cortex *32*, 4447–4463. https://doi.org/10.1093/cercor/bhab494.

51. Kucyi, A., Hove, M.J., Esterman, M., Hutchison, R.M., and Valera, E.M. (2017). Dynamic brain network correlates of spontaneous fluctuations in attention. Cereb. Cortex *27*, 1831–1840. https://doi.org/10.1093/cercor/bhw029.

52. Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn. Sci. *14*, 119–130. https://doi.org/10.1016/j.tics.2010.01.003.

53. Fiser, J., and Lengyel, G. (2019). A common probabilistic framework for perceptual and statistical learning. Curr. Opin. Neurobiol. *58*, 218–228. https://doi.org/10.1016/j.conb.2019.09.007.

54. Santolin, C., and Saffran, J.R. (2018). Constraints on statistical learning across species. Trends Cogn. Sci. *22*, 52–63. https://doi.org/10.1016/J.TICS.2017.10.003.

55. Poldrack, R.A., and Packard, M.G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. Neuropsychologia *41*, 245–251. https://doi.org/10.1016/S0028-3932(02)00157-4.

56. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711. https://doi.org/10.1038/nn1560.

57. Nemeth, D., Janacsek, K., and Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. Front. Comput. Neurosci. *7*, 147. https://doi.org/10.3389/fncom.2013.00147.

58. Smalle, E.H.M., Daikoku, T., Szmalec, A., Duyck, W., and Möttönen, R. (2022). Unlocking adults' implicit statistical learning by cognitive depletion. Proc. Natl. Acad. Sci. USA *119*, e2026011119. https://doi.org/10.1073/pnas.2026011119.

59. Tóth, B., Janacsek, K., Takács, Á., Kóbor, A., Zavecz, Z., and Nemeth, D. (2017). Dynamics of EEG functional connectivity during statistical learning. Neurobiol. Learn. Mem. *144*, 216–229. https://doi.org/10.1016/j.nlm.2017.07.015.

60. Ambrus, G.G., Vékony, T., Janacsek, K., Trimborn, A.B., Kovács, G., and Nemeth, D. (2020). When less is more: Enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. J. Memory Lang. *114*, 104144. https://doi.org/10.1016/j.jml.2020.104144.

61. Baddeley, A., Emslie, H., Kolodny, J., and Duncan, J. (1998). Random generation and the executive control of working memory. Q. J. Exp. Psychol. *51*, 819–852. https://doi.org/10.1080/713755788.

62. Alexandersen, A., Csifcsák, G., Groot, J., and Mittner, M. (2022). The effect of transcranial direct current stimulation on the interplay between executive control, behavioral variability and mind wandering: A registered report. Neuroimage Rep. *2*, 100109. https://doi.org/10.1016/j.ynirp.2022.100109.

63. Boayue, N.M., Csifcsák, G., Kreis, I.V., Schmidt, C., Finn, I., Hovde Vollsund, A.E., and Mittner, M. (2021). The interplay between executive control, behavioural variability and mind wandering: Insights from a high-definition transcranial direct-current stimulation study. Eur. J. Neurosci. *53*, 1498–1516. https://doi.org/10.1111/ejn.15049.

64. Seli, P., Konishi, M., Risko, E.F., and Smilek, D. (2018). The role of task difficulty in theoretical accounts of mind wandering. Conscious. Cogn. *65*, 255–262. https://doi.org/10.1016/j.concog.2018.08.005.

65. Rummel, J., and Boywitt, C.D. (2014). Controlling the stream of thought: Working memory capacity predicts adjustment of mind-wandering to situational demands. Psychon. Bull. Rev. *21*, 1309–1315. https://doi.org/10.3758/s13423-013-0580-3.

66. Barron, E., Riby, L.M., Greer, J., and Smallwood, J. (2011). Absorbed in thought: The effect of mind wandering on the processing of relevant and irrelevant events. Psychol. Sci. *22*, 596–601. https://doi.org/10.1177/0956797611404083.
67. Kóbor, A., Horváth, K., Kardos, Z., Takács, Á., Janacsek, K., Csépe, V., and Nemeth, D. (2019). Tracking the implicit acquisition of nonadjacent transitional probabilities by ERPs. Memory Cogn. *47*, 1546–1566. https://doi.org/10.3758/s13421-019-00949-x.
68. Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., and Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. Biol. Psychol. *135*, 180–193. https://doi.org/10.1016/j.biopsycho.2018.04.001.
69. Jubera-Garcia, E., Gevers, W., and Van Opstal, F. (2021). Local build-up of sleep pressure could trigger mind wandering: Evidence from sleep, circadian and mind wandering research. Biochem. Pharmacol. *191*, 114478. https://doi.org/10.1016/j.bcp.2021.114478.
70. Jubera-García, E., Vermeylen, L., Peigneux, P., Gevers, W., and Opstal, F.V. (2021). Local use-dependent activity triggers mind wandering: Resource depletion or executive dysfunction? J. Exper. Psychol. Human Percept. Perform. *47*, 1575. https://doi.org/10.1037/xhp0000959.
71. Wienke, C., Bartsch, M.V., Vogelgesang, L., Reichert, C., Hinrichs, H., Heinze, H.-J., and Dürschmid, S. (2021). Mind-wandering is accompanied by both local sleep and enhanced processes of spatial attention allocation. Cereb. Cortex Commun. *2*, tgab001. https://doi.org/10.1093/texcom/tgab001.
72. Rasch, B., and Born, J. (2013). About sleep's role in memory. Physiol. Rev. *93*, 681–766. https://doi.org/10.1152/physrev.00032.2012.
73. Diekelmann, S., and Born, J. (2010). The memory function of sleep. Nat. Rev. Neurosci. *11*, 114–126. https://doi.org/10.1038/nrn2762.
74. Krueger, J.M., Nguyen, J.T., Dykstra-Aiello, C.J., and Taishi, P. (2019). Local sleep. Sleep Med. Rev. *43*, 14–21. https://doi.org/10.1016/j.smrv.2018.10.001.
75. Vyazovskiy, V.V., Olcese, U., Hanlon, E.C., Nir, Y., Cirelli, C., and Tononi, G. (2011). Local sleep in awake rats. Nature *472*, 443–447. https://doi.org/10.1038/nature10009.
76. Wamsley, E.J., and Summer, T. (2020). Spontaneous entry into an "offline" state during wakefulness: A mechanism of memory consolidation? J. Cogn. Neurosci. *32*, 1714–1734. https://doi.org/10.1162/jocn_a_01587.
77. Pan, S.C., and Rickard, T.C. (2015). Sleep and motor learning: Is there room for consolidation? Psychol. Bull. *141*, 812–834. https://doi.org/10.1037/BUL0000009.
78. Vékony, T., Ambrus, G.G., Janacsek, K., and Nemeth, D. (2022). Cautious or causal? Key implicit sequence learning paradigms should not be overlooked when assessing the role of DLPFC (Commentary on Prutean et al.). Cortex *148*, 222. https://doi.org/10.1016/j.cortex.2021.10.001.
79. Blondé, P., Girardeau, J.C., Sperduti, M., and Piolino, P. (2022). A wandering mind is a forgetful mind: A systematic review on the influence of mind wandering on episodic memory encoding. Neurosci. Biobehav. Rev. *132*, 774–792. https://doi.org/10.1016/j.neubiorev.2021.11.015.
80. Maillet, D., and Schacter, D.L. (2016). When the mind wanders: Distinguishing stimulus-dependent from stimulus-independent thoughts during incidental encoding in young and older adults. Psychol. Aging *31*, 370–379. https://doi.org/10.1037/pag0000099.
81. Maillet, D., Seli, P., and Schacter, D.L. (2017). Mind-wandering and task stimuli: Stimulus-dependent thoughts influence performance on memory tasks and are more often past- versus future-oriented. Conscious. Cogn. *52*, 55–67. https://doi.org/10.1016/j.concog.2017.04.014.
82. Spencer, M., Kaschak, M.P., Jones, J.L., and Lonigan, C.J. (2015). Statistical learning is related to early literacy-related skills. Read. Writ. *28*, 467–490. https://doi.org/10.1007/S11145-014-9533-0/METRICS.
83. Romano Bergstrom, J.C., Howard, J.H., and Howard, D.V. (2012). Enhanced implicit sequence learning in college-age video game players and musicians. Appl. Cogn. Psychol. *26*, 91–96. https://doi.org/10.1002/acp.1800.
84. Lieberman, M.D. (2000). Intuition: A social cognitive neuroscience approach. Psychol. Bull. *126*, 109–137. https://doi.org/10.1037/0033-2909.126.1.109.
85. Nemeth, D., Janacsek, K., Csifcsak, G., Szvoboda, G., Howard, J.H., and Howard, D.V. (2011). Interference between sentence processing and probabilistic implicit sequence learning. PLoS One *6*, e17577.
86. Szücs-Bencze, L., Fanuel, L., Szabó, N., Quentin, R., Nemeth, D., and Vékony, T. (2023). Manipulating the Rapid Consolidation Periods in a Learning Task Affects General Skills More than Statistical Learning and Changes the Dynamics of Learning. eNeuro *10*, ENEURO.0228-22.2022. https://doi.org/10.1523/ENEURO.0228-22.2022.
87. de Leeuw, J.R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. Behav. Res. Methods *47*, 1–12. https://doi.org/10.3758/s13428-014-0458-y.
88. Vékony, T. (2021). Alternating Serial Reaction Time Task created with jsPsych. Preprint at. https://doi.org/10.5281/ZENODO.7124730.
89. Ward, A.F., and Wegner, D.M. (2013). Mind-blanking: When the mind goes away. Front. Psychol. *4*, 650. https://doi.org/10.3389/fpsyg.2013.00650.
90. Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N., and Evershed, J.K. (2020). Gorilla in our midst: An online behavioral experiment builder. Behav. Res. Methods *52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x.
91. Howard, D.V., Howard, J.H., Japikse, K., DiYanni, C., Thompson, A., and Somberg, R. (2004). Implicit sequence learning: Effects of level of structure, adult age, and extended practice. Psychol. Aging *19*, 79–92. https://doi.org/10.1037/0882-7974.19.1.79.
92. Singmann, H., Bolker, B., Westfall, J., and Aust, F. (2016). afex: Analysis of Factorial Experiments. R package version 0.16-1.e.
93. Tiedemann, F. (2020). ggpol: Visualizing social science data with 'ggplot2'. R package version 0.0, 7.
94. Wilke, C.O., Wickham, H., and Wilke, M.C.O. (2019). Package 'cowplot'. Streamlined plot theme and plot annotations for 'ggplot2'.
95. Wickham, H. (2011). ggplot2. WIREs Computational Stats. *3*, 180–185. https://doi.org/10.1002/wics.147.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw data	This paper, OSF	OSF: https://osf.io/t4ab3/
Software and algorithms		
Experimental task code	This paper, OSF	OSF: https://osf.io/t4ab3/
Analysis code	This paper, OSF	OSF: https://osf.io/t4ab3/
Preview of the experimental task	This Paper, Gorilla Open Materials	Gorilla Open Materials: https://app.gorilla.sc/openmaterials/581806

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

A total of 152 participants completed an online experiment. The sample consisted of students from Eötvös Loránd University in Budapest, Hungary, who received course credit for their participation. Quality control measures were implemented to ensure the integrity of the data, resulting in the exclusion of 17 participants for non-compliance with instructions (see section "quality control of the data"). The final sample contained 135 participants (114 females; M_{age} = 22.58 years ± 5.67 SD; 103 Bachelor and 32 Master students), each of them allocated to the experimental group. This study did not include analyses based on age, sex, gender, ancestry, race and ethnicity, which may limit the generalizability of our results. Informed consent was obtained from all participants, and the study was approved by the Research Ethics Committee of Eötvös Loránd University, Budapest, Hungary. The sample size was determined based on a previous online study where within-subject differences were reported with the ASRT task.[86]

METHOD DETAILS

Alternating Serial Reaction Time task

The ASRT task was programmed in JavaScript using the jsPsych library v.6.1.0.[87,88] The task involved presenting participants with a visual stimulus (a drawing of a dog's head) in one of four horizontal locations on the screen, and participants were instructed to indicate the location of the target stimulus by pressing the corresponding key on the keyboard (S, F, J, or L keys from left to right) (Figure 1A). In case of correct response, the target stimulus disappeared, and after a 120 ms interstimulus interval, the next stimulus appeared. In case of an incorrect response, the target stimulus remained in place until the first correct response. The stimuli followed a probabilistic eight-element sequence, with pattern and random elements alternating with each other (e.g., 2 – r – 4 – r – 3 – r – 1 – r, where *r* indicates a random location, and the numbers represent the predetermined positions from left to the right). Each participant was assigned to one of 24 possible sequences, which they were exposed to throughout the task. The ASRT task was composed of 25 blocks, with each block containing ten repetitions of the eight-element sequence. After each block, participants had to take a short break and were instructed to answer the thought probes before continuing (Figure 1B).

The ASRT task featured a probabilistic sequence structure where certain runs of three consecutive stimuli (triplets) appeared with a higher probability (high-probability triplets) than others (low-probability triplets). A *trial* refers to a single element in the sequence that could be either a pattern or random element, and, crucially, also the last element in a high- or low-probability triplet (high-vs. low-probability trials). It is important to emphasize that the analysis hinges on whether the provided trial constitutes the final element of a high- or low-probability triplet, rather than its classification as a pattern or a random element within the alternating sequence. For example, in a sequence such as 2 – r – 4 – r – 3 – r – 1 – r, triplets such as 2-X-4, 4-X-3, 3-X-1, and 1-X-2 (where X represents the middle element of a triplet) occurred more frequently than triplets such as 2-X-1 or 2-X-3 (Figure 1C). Please note that only three triplets are highlighted on Figure 1C for visualization purposes [2(P)-1(R)-4(P) as a pattern-ending high-probability triplet, 2(R)-3(P)-4(R) as a random-ending high-probability triplet, and 4(R)-1(P)-2(R) as a random-ending low-probability triplet]. However, every three consecutive elements form either a high- or low-probability triplet. Therefore, in the above example, from these 8 consecutive elements - 2(P)-1(R)-4(P)-2(R)-3(P)-4(R)-1(P)-2(R) - 6 triplets can be formed. If we consider the pattern element 2 as a starting point, then the triplets are in the following order: 2(P)-1(R)-4(P), 1(R)-4(P)-2(R), 4(P)-2(R)-3(P), 2(R)-3(P)-4(R), 3(P)-4(R)-1(P) (these are all high-probability triplets), and 4(R)-1(P)-2(R) (low-probability triplet). When referring to triplet type in the later parts, the focus is on trials that serve as the final element of a high- or low-probability triplet.

Throughout the task, a total of 64 distinct triplets could potentially occur (16 with high-probability and 48 with low-probability). High-probability triplets could be formed by either having two pattern trials and one random trial in the center (occurring in 50% of trials) or by having two random trials and one pattern trial in the center (occurring in 12.5% of trials). Of all trials, 62.5% represented the last element of a high-probability triplet (referred to as high-probability trial), and 37.5% were assigned to the last element of a low-probability triplet (referred to as low-probability trial) (Figure 1D).

Thought probes

After each block of the ASRT task, participants were asked to reflect on their thoughts and respond to questions about their MW state. Even though the current study primarily investigated if MW, defined more generally as task-unrelated thought, is associated with implicit statistical learning, in a more exploratory approach, we also tested if such beneficial effects are linked to 1) MW episodes with reportable content in contrast to "mind blanking" (MB),[89] and to 2) spontaneous or deliberate manifestations of MW.

Thus, after each block, participants were asked to respond to three questions aimed at distinguishing between 1) on-task vs. off-task (MW) periods, 2) MW and MB periods, and 3) deliberate vs. non-deliberate/spontaneous episodes. The first question (Q1) was "*To what degree were you focusing on the task just before this question?"* (1 - Not at all; 4 - Completely). Here, a response of 1 meant that their thoughts were completely diverted (e.g., friends, weekend plans), and a response of 4 meant that they were only thinking about the task at hand (where the dog's head appears, and which key to press as soon as possible). The second question (Q2) was "*To the degree to which you were not focusing on the task, were you thinking of something in particular or just thinking about nothing?*" (1 - I was thinking about nothing; 4 - I was thinking about something in particular). Here, the response of 1 meant that their mind wandered away from the task, but they were not thinking about anything. The response of 4 meant that they were thinking about something while engaging in mind wandering (e.g., a book, recent events, the task was too easy/difficult, being hungry, it was uncomfortable to sit and do the task, etc.). The third question (Q3) was "*Were you deliberate about where you focused your attention (either on-task or elsewhere) or did it happen spontaneously?*" (1 - I was completely spontaneous; 4 - I was completely deliberate). Here, a response of 1 meant that maintaining their attentional focus occurred effortlessly, without deliberation. The response of 4 meant that they consciously directed their attention somewhere. The first question (Q1) has been used in previous studies,[49,62] whereas the other two were phrased to either distinguish between MW with reportable content vs. MB (Q2) or between spontaneous vs. deliberate MW (Q3). Even though it is common to directly differentiate between on-task periods and either MW vs. MB, or unintentional vs. intentional MW in a single thought probe,[7,64] we decided to explore these dimensions of MW in two follow-up questions in order to avoid presenting participants with too many response options at once, and also, to gain a more nuanced view about their mental states during the ASRT task. Participants were asked to select their answers by clicking the corresponding checkbox using their mouse or touchpad. For the results pertaining to Q2 and Q3, see Figures S1–S4 and Tables S1–S4, respectively.

Procedures

The Gorilla Experiment Builder (https://www.gorilla.sc) was utilized to host the experiment.[90] At the beginning of the study, participants completed a picture description task to test their compliance. Then, they were presented with instructions for the ASRT task, which involved pressing the key corresponding to the location of the target stimulus as quickly and accurately as possible, using their left middle and index fingers and their right index and middle fingers from left to right, respectively. Participants were informed that after each block of the ASRT task, they would be asked three questions to evaluate their thoughts in the previous block and assess their level of MW which was operationalized as task-unrelated perceptions, thoughts, or memories. A detailed explanation was provided on the different options along with examples of how participants should respond in various scenarios. Subsequently, participants completed a short quiz to evaluate their understanding of how to answer the questions about their thoughts, with feedback and explanations provided afterward. Participants had the option to retake the quiz or proceed to the task.

Following the two initial practice blocks of the ASRT task with random stimuli, participants completed 25 additional blocks of the ASRT task. After responding to the thought probes in each block, participants received performance feedback, which included information on both mean speed and accuracy. In order to guarantee that the ASRT task functioned in a similar manner to previous studies,[78] we assessed the participants' conscious knowledge of the hidden sequence. Participants were asked a series of questions at the end of the ASRT task. Specifically, they were asked if they noticed anything unusual or any regularities in the task, and if so, to elaborate on their response. None of the participants were able to accurately describe the alternating sequence.

After completing the ASRT task, participants were asked to respond to demographic questions (age, gender, education, etc.). Additionally, they were asked to complete a short questionnaire about their surroundings during the online experiment to test potential non-compliance with instructions during task completion.

Quality control of the data

To ensure the validity of the study's conclusions, pre-registered exclusion criteria were employed to remove participants who did not comply with the instructions or failed the attention tasks. Two participants were excluded based on their performance on the picture description task, while five participants gave incorrect answers to out-of-context questions in the questionnaires (e.g., select a specific response or write a specific answer to an open-ended question). Four participants reported that they had restarted the ASRT task, indicating severe non-compliance with the instructions. One participant was excluded based on evidence from the ASRT task data for restarting the task at one point. Additionally, six participants were excluded for unusually low performance on the ASRT task (<80% accuracy), and one participant was excluded for random button pressing in several blocks. A total of 135 participants remained in the final sample (please note that some participants met more than one exclusion criterion).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical learning scores

We used Python v3.9 for processing the responses in the ASRT task. Each trial was categorized based on the two preceding trials as the last element of a high- or low-probability triplet. To minimize the impact of pre-existing response tendencies,[91] we excluded trials such as trills (e.g., 1-2-1) and, repetitions (e.g., 2-2-2). The first two trials of each block were also removed, as they could not be categorized as the third element of a triplet, and trials with a reaction time falling outside of ±1.5 times the individual interquartile range were also removed from the analysis. We defined two types of learning: statistical learning and visuomotor performance. *Statistical learning* was operationalized as the difference in accuracy between high-probability and low-probability trials (i.e., between the third element of a high-probability triplet and the third element of a low-probability triplet, henceforth referred to as *statistical learning score*). *Visuomotor performance*, on the other hand, was operationalized as the overall accuracy and reaction time performance on the task and their changes over time regardless of item probability.

Calculation of MW scores from thought probe responses

A dichotomous variable was created by categorizing the answers to thought probe Q1 into MW (off-task) (answers 1–2) and on-task (answers 3–4) periods. Contrasts between MW with reportable content vs. MB were created by dichotomizing responses to Q2 (1–2 vs. 3–4) exclusively to those thought probes where Q1 reports indicated off-task periods. In a similar vein, we contrasted task performance between spontaneous vs. deliberate MW by categorizing responses to Q3 (1–2 vs. 3–4) for thought probes indicating MW for Q1. Additionally, the mean MW scores across all thought probes were computed for each participant, enabling between-subject comparison via median split (higher MW propensity: M = 2.483 ± 0.347 SD, lower MW propensity: M = 3.428 ± 0.276 SD). For the results of questions Q2 and Q3, please refer to Figures S1–S4 and Tables S1–S4, respectively. For the results of the between-subject comparison, please refer to Tables S7–S10.

Statistical analysis

The statistical analysis was performed using *R* 4.2.3. Simple regressions were fitted with the *lm* function of the *lme4* package, and linear mixed models were fitted with the *mixed* function from the *afex* package[92] with sum-to-zero contrasts. For each linear mixed model, we first fitted the maximal random-effects structure (i.e., including random effects for all variables and allowing correlations between them), and then reduced it to achieve convergence by removing correlations between random slopes or the random slopes themselves. The numerical fixed factors were mean-centered. Estimated marginal means were computed with the *emmeans* R package. Alpha level 0.05 was applied to all analyses. Figures were created with the *ggplot2* supplemented by the *cowplot, ggpol*, and *afex* R packages.[92–95]

To investigate how MW changed throughout the learning process, we performed simple regressions using the mean MW scores and the proportion of participants engaged in MW, with Block serving as the predictor. To explore how MW impacted statistical learning and visuomotor performance, we computed the median reaction times and mean accuracy of the ASRT task for each participant in every block. These measures were used as outcome variables in linear mixed models with the Block (Block 1–25), Triplet Type (last element of a high- vs. low-probability triplet), and MW (MW vs. on-task periods) as well as their interactions as fixed effects in the models. For reaction times, the final model achieved convergence using the participants as random intercepts with by-participant uncorrelated slopes for the Block and MW factors. For accuracy, the final model achieved convergence using the participants as random intercepts with by-participant uncorrelated slopes for Triplet Type, Block and MW factors. Please note that the Triplet Type main effect and interactions involving this factor indicate differences in statistical learning, while main effects and interactions without it can be interpreted as differences in visuomotor performance. For further pre-registered analyses, see Figures S8, S9, and Tables S17–S21.

ADDITIONAL RESOURCES

The study was pre-registered on OSF, which is available at the following link, OSF: https://osf.io/cq6pg. This study is not conducted within the framework of a clinical trial.